SEQUENCE LISTING

<110> DSM IP Assets B.V.

<120> Feedback-resistant mevalonate kinases

<130> Case 21788

<150> EP 03012294.9
<151> 2003-06-12

<160> 33

<170> PatentIn version 3.2

<210> 1
<211> 378
<212> PRT
<213> Paracoccus zeaxanthinifaciens

<400> 1

Met Ser Thr Gly Arg Pro Glu Ala Gly Ala His Ala Pro Gly Lys Leu
1 5 10 15

Ile Leu Ser Gly Glu His Ser Val Leu Tyr Gly Ala Pro Ala Leu Ala
20 25 30

Met Ala Ile Ala Arg Tyr Thr Glu Val Trp Phe Thr Pro Leu Gly Ile
35 40 45

Gly Glu Gly Ile Arg Thr Thr Phe Ala Asn Leu Ser Gly Gly Ala Thr
50 55 60

Tyr Ser Leu Lys Leu Leu Ser Gly Phe Lys Ser Arg Leu Asp Arg Arg
65 70 75 80

Phe Glu Gln Phe Leu Asn Gly Asp Leu Lys Val His Lys Val Leu Thr
85 90 95

His Pro Asp Asp Leu Ala Val Tyr Ala Leu Ala Ser Leu Leu His Asp
100 105 110

Lys Pro Pro Gly Thr Ala Ala Met Pro Gly Ile Gly Ala Met His His
115 120 125

Leu Pro Arg Pro Gly Glu Leu Gly Ser Arg Thr Glu Leu Pro Ile Gly
130 135 140

Ala Gly Met Gly Ser Ser Ala Ala Ile Val Ala Ala Thr Thr Val Leu
145 150 155 160

Phe Glu Thr Leu Leu Asp Arg Pro Lys Thr Pro Glu Gln Arg Phe Asp

165 170 175

Arg Val Arg Phe Cys Glu Arg Leu Lys His Gly Lys Ala Gly Pro Ile
180 185 190

Asp Ala Ala Ser Val Val Arg Gly Gly Leu Val Arg Val Gly Gly Asn
195 200 205

Gly Pro Gly Ser Ile Ser Ser Phe Asp Leu Pro Glu Asp His Asp Leu
210 215 220

Val Ala Gly Arg Gly Trp Tyr Trp Val Leu His Gly Arg Pro Val Ser
225 230 235 240

Gly Thr Gly Glu Cys Val Ser Ala Val Ala Ala Ala His Gly Arg Asp
245 250 255

Ala Ala Leu Trp Asp Ala Phe Ala Val Cys Thr Arg Ala Leu Glu Ala
260 265 270

Ala Leu Leu Ser Gly Gly Ser Pro Asp Ala Ala Ile Thr Glu Asn Gln
275 280 285

Arg Leu Leu Glu Arg Ile Gly Val Val Pro Ala Ala Thr Gln Ala Leu
290 295 300

Val Ala Gln Ile Glu Glu Ala Gly Gly Ala Ala Lys Ile Cys Gly Ala
305 310 315 320

Gly Ser Val Arg Gly Asp His Gly Gly Ala Val Leu Val Arg Ile Asp
325 330 335

Asp Ala Gln Ala Met Ala Ser Val Met Ala Arg His Pro Asp Leu Asp
340 345 350

Trp Ala Pro Leu Arg Met Ser Arg Thr Gly Ala Ala Pro Gly Pro Ala
355 360 365

Pro Arg Ala Gln Pro Leu Pro Gly Gln Gly
370 375

<210> 2
<211> 396
<212> PRT
<213> homo sapiens

<400> 2

Met Leu Ser Glu Val Leu Leu Val Ser Ala Pro Gly Lys Val Ile Leu

```
1               5                  10                  15

His Gly Glu His Ala Val Val His Gly Lys Val Ala Leu Ala Val Ser
            20                  25                  30

Leu Asn Leu Arg Thr Phe Leu Arg Leu Gln Pro His Ser Asn Gly Lys
        35                  40                  45

Val Asp Leu Ser Leu Pro Asn Ile Gly Ile Lys Arg Ala Trp Asp Val
    50                  55                  60

Ala Arg Leu Gln Ser Leu Asp Thr Ser Phe Leu Glu Gln Gly Asp Val
65                  70                  75                  80

Thr Thr Pro Thr Ser Glu Gln Val Glu Lys Leu Lys Glu Val Ala Gly
                85                  90                  95

Leu Pro Asp Asp Cys Ala Val Thr Glu Arg Leu Ala Val Leu Ala Phe
            100                 105                 110

Leu Tyr Leu Tyr Leu Ser Ile Cys Arg Lys Gln Arg Ala Leu Pro Ser
        115                 120                 125

Leu Asp Ile Val Val Trp Ser Glu Leu Pro Pro Gly Ala Gly Leu Gly
    130                 135                 140

Ser Ser Ala Ala Tyr Ser Val Cys Leu Ala Ala Ala Leu Leu Thr Val
145                 150                 155                 160

Cys Glu Glu Ile Pro Asn Pro Leu Lys Asp Gly Asp Cys Val Asn Arg
                165                 170                 175

Trp Thr Lys Glu Asp Leu Glu Leu Ile Asn Lys Trp Ala Phe Gln Gly
            180                 185                 190

Glu Arg Met Ile His Gly Asn Pro Ser Gly Val Asp Asn Ala Val Ser
        195                 200                 205

Thr Trp Gly Gly Ala Leu Arg Tyr His Gln Gly Lys Ile Ser Ser Leu
    210                 215                 220

Lys Arg Ser Pro Ala Leu Gln Ile Leu Leu Thr Asn Thr Lys Val Pro
225                 230                 235                 240

Arg Asn Thr Arg Ala Leu Val Ala Gly Val Arg Asn Arg Leu Leu Lys
                245                 250                 255
```

```
Phe Pro Glu Ile Val Ala Pro Leu Leu Thr Ser Ile Asp Ala Ile Ser
            260                 265                 270

Leu Glu Cys Glu Arg Val Leu Gly Glu Met Gly Glu Ala Pro Ala Pro
        275                 280                 285

Glu Gln Tyr Leu Val Leu Glu Glu Leu Ile Asp Met Asn Gln His His
    290                 295                 300

Leu Asn Ala Leu Gly Val Gly His Ala Ser Leu Asp Gln Leu Cys Gln
305                 310                 315                 320

Val Thr Arg Ala Arg Gly Leu His Ser Lys Leu Thr Gly Ala Gly Gly
                325                 330                 335

Gly Gly Cys Gly Ile Thr Leu Leu Lys Pro Gly Leu Glu Gln Pro Glu
            340                 345                 350

Val Glu Ala Thr Lys Gln Ala Leu Thr Ser Cys Gly Phe Asp Cys Leu
        355                 360                 365

Glu Thr Ser Ile Gly Ala Pro Gly Val Ser Ile His Ser Ala Thr Ser
    370                 375                 380

Leu Asp Ser Arg Val Gln Gln Ala Leu Asp Gly Leu
385                 390                 395
```

<210> 3
<211> 395
<212> PRT
<213> mouse

<400> 3

```
Met Leu Ser Glu Ala Leu Leu Val Ser Ala Pro Gly Lys Val Ile Leu
1               5                   10                  15

His Gly Glu His Ala Val Val His Gly Lys Val Ala Leu Ala Ala Ala
            20                  25                  30

Leu Asn Leu Arg Thr Phe Leu Leu Leu Arg Pro Gln Ser Asn Gly Lys
        35                  40                  45

Val Ser Val Asn Leu Pro Asn Ile Gly Ile Lys Gln Val Trp Asp Val
    50                  55                  60

Gly Met Leu Gln Arg Leu Asp Thr Ser Phe Leu Glu Gln Gly Asp Val
65                  70                  75                  80
```

```
Ser Val Pro Thr Leu Glu Gln Leu Glu Lys Leu Lys Lys Met Gly Asp
                85                  90                  95

Leu Pro Arg Asp Arg Ala Gly Asn Glu Gly Met Ala Leu Leu Ala Phe
            100                 105                 110

Leu Tyr Leu Tyr Leu Ala Ile Cys Arg Lys Gln Arg Thr Leu Pro Ser
        115                 120                 125

Leu Asp Met Val Val Trp Ser Glu Leu Pro Pro Gly Ala Gly Leu Gly
    130                 135                 140

Ser Ser Ala Ala Tyr Ser Val Cys Leu Ala Ala Ala Leu Leu Thr Ala
145                 150                 155                 160

Cys Glu Glu Val Ser Asn Pro Leu Lys Asp Gly Val Ser Val Ser Arg
                165                 170                 175

Trp Pro Glu Glu Asp Leu Lys Ser Ile Asn Lys Trp Ala Phe Glu Gly
            180                 185                 190

Glu Arg Val Ile His Gly Asn Pro Ser Gly Val Asp Asn Ala Val Ser
        195                 200                 205

Thr Trp Gly Gly Ala Leu Arg Phe Gln Gln Gly Thr Met Ser Ser Leu
    210                 215                 220

Lys Ser Leu Pro Ser Leu Gln Ile Leu Leu Thr Asn Thr Lys Val Pro
225                 230                 235                 240

Arg Ser Thr Lys Ala Leu Val Ala Ala Val Arg Ser Arg Leu Thr Lys
                245                 250                 255

Phe Pro Glu Ile Val Ala Pro Leu Leu Thr Ser Ile Asp Ala Ile Ser
            260                 265                 270

Leu Glu Cys Glu Arg Val Leu Gly Glu Met Val Ala Ala Pro Val Pro
        275                 280                 285

Glu Gln Tyr Leu Val Leu Glu Glu Leu Ile Asp Met Asn Gln His His
    290                 295                 300

Leu Asn Ala Leu Gly Val Gly His Asn Ser Leu Asp Gln Leu Cys Gln
305                 310                 315                 320

Val Thr Ala Ala His Gly Leu His Ser Lys Leu Thr Gly Ala Gly Gly
                325                 330                 335
```

```
Gly Gly Cys Gly Ile Thr Leu Leu Lys Pro Gly Leu Glu Gln Ala Thr
            340                 345                 350

Val Glu Ala Ala Lys Gln Ala Leu Thr Ser Cys Gly Phe Asp Cys Trp
        355                 360                 365

Glu Thr Ser Ile Gly Ala Pro Gly Val Ser Thr His Ser Ala Ala Ala
    370                 375                 380

Val Gly Asp Pro Val Arg Gln Ala Leu Gly Leu
385                 390                 395
```

<210> 4
<211> 395
<212> PRT
<213> rat

<400> 4

```
Met Leu Ser Glu Val Leu Leu Val Ser Ala Pro Gly Lys Val Ile Leu
1               5                   10                  15

His Gly Glu His Ala Val Val His Gly Lys Val Ala Leu Ala Val Ala
            20                  25                  30

Leu Asn Leu Arg Thr Phe Leu Val Leu Arg Pro Gln Ser Asn Gly Lys
        35                  40                  45

Val Ser Leu Asn Leu Pro Asn Val Gly Ile Lys Gln Val Trp Asp Val
    50                  55                  60

Ala Thr Leu Gln Leu Leu Asp Thr Gly Phe Leu Glu Gln Gly Asp Val
65                  70                  75                  80

Pro Ala Pro Thr Leu Glu Gln Leu Glu Lys Leu Lys Lys Val Ala Gly
                85                  90                  95

Leu Pro Arg Asp Cys Val Gly Asn Glu Gly Leu Ser Leu Leu Ala Phe
            100                 105                 110

Leu Tyr Leu Tyr Leu Ala Ile Cys Arg Lys Gln Arg Thr Leu Pro Ser
        115                 120                 125

Leu Asp Ile Met Val Trp Ser Glu Leu Pro Pro Gly Ala Gly Leu Gly
    130                 135                 140

Ser Ser Ala Ala Tyr Ser Val Cys Val Ala Ala Ala Leu Leu Thr Ala
145                 150                 155                 160
```

```
Cys Glu Glu Val Thr Asn Pro Leu Lys Asp Arg Gly Ser Ile Gly Ser
                165                 170                 175

Trp Pro Glu Glu Asp Leu Lys Ser Ile Asn Lys Trp Ala Tyr Glu Gly
            180                 185                 190

Glu Arg Val Ile His Gly Asn Pro Ser Gly Val Asp Asn Ser Val Ser
        195                 200                 205

Thr Trp Gly Gly Ala Leu Arg Tyr Gln Gln Gly Lys Met Ser Ser Leu
    210                 215                 220

Lys Arg Leu Pro Ala Leu Gln Ile Leu Leu Thr Asn Thr Lys Val Pro
225                 230                 235                 240

Arg Ser Thr Lys Ala Leu Val Ala Gly Val Arg Ser Arg Leu Ile Lys
                245                 250                 255

Phe Pro Glu Ile Met Ala Pro Leu Leu Thr Ser Ile Asp Ala Ile Ser
            260                 265                 270

Leu Glu Cys Glu Arg Val Leu Gly Glu Met Ala Ala Ala Pro Val Pro
        275                 280                 285

Glu Gln Tyr Leu Val Leu Glu Glu Leu Met Asp Met Asn Gln His His
    290                 295                 300

Leu Asn Ala Leu Gly Val Gly His Ala Ser Leu Asp Gln Leu Cys Gln
305                 310                 315                 320

Val Thr Ala Ala His Gly Leu His Ser Lys Leu Thr Gly Ala Gly Gly
                325                 330                 335

Gly Gly Cys Gly Ile Thr Leu Leu Lys Pro Gly Leu Glu Arg Ala Lys
            340                 345                 350

Val Glu Ala Ala Lys Gln Ala Leu Thr Gly Cys Gly Phe Asp Cys Trp
        355                 360                 365

Glu Thr Ser Ile Gly Ala Pro Gly Val Ser Met His Ser Ala Thr Ser
    370                 375                 380

Ile Glu Asp Pro Val Arg Gln Ala Leu Gly Leu
385                 390                 395
```

<210> 5
<211> 378

<212> PRT
<213> Arabidopsis thaliana

<400> 5

```
Met Glu Val Lys Ala Arg Ala Pro Gly Lys Ile Ile Leu Ala Gly Glu
1               5                  10                  15

His Ala Val Val His Gly Ser Thr Ala Val Ala Ala Ala Ile Asp Leu
            20                  25                  30

Tyr Thr Tyr Val Thr Leu Arg Phe Pro Leu Pro Ser Ala Glu Asn Asn
        35                  40                  45

Asp Arg Leu Thr Leu Gln Leu Lys Asp Ile Ser Leu Glu Phe Ser Trp
    50                  55                  60

Ser Leu Ala Arg Ile Lys Glu Ala Ile Pro Tyr Asp Ser Ser Thr Leu
65                  70                  75                  80

Cys Arg Ser Thr Pro Ala Ser Cys Ser Glu Glu Thr Leu Lys Ser Ile
                85                  90                  95

Ala Val Leu Val Glu Glu Gln Asn Leu Pro Lys Glu Lys Met Trp Leu
            100                 105                 110

Ser Ser Gly Ile Ser Thr Phe Leu Trp Leu Tyr Thr Arg Ile Ile Gly
        115                 120                 125

Phe Asn Pro Ala Thr Val Val Ile Asn Ser Glu Leu Pro Tyr Gly Ser
    130                 135                 140

Gly Leu Gly Ser Ser Ala Ala Leu Cys Val Ala Leu Thr Ala Ala Leu
145                 150                 155                 160

Leu Ala Ser Ser Ile Ser Glu Lys Thr Arg Gly Asn Gly Trp Ser Ser
                165                 170                 175

Leu Asp Glu Thr Asn Leu Glu Leu Leu Asn Lys Trp Ala Phe Glu Gly
            180                 185                 190

Glu Lys Ile Ile His Gly Lys Pro Ser Gly Ile Asp Asn Thr Val Ser
        195                 200                 205

Ala Tyr Gly Asn Met Ile Lys Phe Cys Ser Gly Glu Ile Thr Arg Leu
    210                 215                 220

Gln Ser Asn Met Pro Leu Arg Met Leu Ile Thr Asn Thr Arg Val Gly
225                 230                 235                 240
```

```
Arg Asn Thr Lys Ala Leu Val Ser Gly Val Ser Gln Arg Ala Val Arg
                245                 250                 255

His Pro Asp Ala Met Lys Ser Val Phe Asn Ala Val Asp Ser Ile Ser
            260                 265                 270

Lys Glu Leu Ala Ala Ile Ile Gln Ser Lys Asp Glu Thr Ser Val Thr
        275                 280                 285

Glu Lys Glu Glu Arg Ile Lys Glu Leu Met Glu Met Asn Gln Gly Leu
    290                 295                 300

Leu Leu Ser Met Gly Val Ser His Ser Ser Ile Glu Ala Val Ile Leu
305                 310                 315                 320

Thr Thr Val Lys His Lys Leu Val Ser Lys Leu Thr Gly Ala Gly Gly
                325                 330                 335

Gly Gly Cys Val Leu Thr Leu Leu Pro Thr Gly Thr Val Val Asp Lys
            340                 345                 350

Val Val Glu Glu Leu Glu Ser Ser Gly Phe Gln Cys Phe Thr Ala Leu
        355                 360                 365

Ile Gly Gly Asn Gly Ala Gln Ile Cys Tyr
    370                 375
```

<210> 6
<211> 443
<212> PRT
<213> Saccharomyces cerevisiae

<400> 6

```
Met Ser Leu Pro Phe Leu Thr Ser Ala Pro Gly Lys Val Ile Ile Phe
1               5                   10                  15

Gly Glu His Ser Ala Val Tyr Asn Lys Pro Ala Val Ala Ala Ser Val
            20                  25                  30

Ser Ala Leu Arg Thr Tyr Leu Leu Ile Ser Glu Ser Ser Ala Pro Asp
        35                  40                  45

Thr Ile Glu Leu Asp Phe Pro Asp Ile Ser Phe Asn His Lys Trp Ser
    50                  55                  60

Ile Asn Asp Phe Asn Ala Ile Thr Glu Asp Gln Val Asn Ser Gln Lys
65                  70                  75                  80
```

```
Leu Ala Lys Ala Gln Gln Ala Thr Asp Gly Leu Ser Gln Glu Leu Val
                85                  90                  95

Ser Leu Leu Asp Pro Leu Leu Ala Gln Leu Ser Glu Ser Phe His Tyr
            100                 105                 110

His Ala Ala Phe Cys Phe Leu Tyr Met Phe Val Cys Leu Cys Pro His
        115                 120                 125

Ala Lys Asn Ile Lys Phe Ser Leu Lys Ser Thr Leu Pro Ile Gly Ala
    130                 135                 140

Gly Leu Gly Ser Ser Ala Ser Ile Ser Val Ser Leu Ala Leu Ala Met
145                 150                 155                 160

Ala Tyr Leu Gly Gly Leu Ile Gly Ser Asn Asp Leu Glu Lys Leu Ser
                165                 170                 175

Glu Asn Asp Lys His Ile Val Asn Gln Trp Ala Phe Ile Gly Glu Lys
            180                 185                 190

Cys Ile His Gly Thr Pro Ser Gly Ile Asp Asn Ala Val Ala Thr Tyr
        195                 200                 205

Gly Asn Ala Leu Leu Phe Glu Lys Asp Ser His Asn Gly Thr Ile Asn
    210                 215                 220

Thr Asn Asn Phe Lys Phe Leu Asp Asp Phe Pro Ala Ile Pro Met Ile
225                 230                 235                 240

Leu Thr Tyr Thr Arg Ile Pro Arg Ser Thr Lys Asp Leu Val Ala Arg
                245                 250                 255

Val Arg Val Leu Val Thr Glu Lys Phe Pro Glu Val Met Lys Pro Ile
            260                 265                 270

Leu Asp Ala Met Gly Glu Cys Ala Leu Gln Gly Leu Glu Ile Met Thr
        275                 280                 285

Lys Leu Ser Lys Cys Lys Gly Thr Asp Asp Glu Ala Val Glu Thr Asn
    290                 295                 300

Asn Glu Leu Tyr Glu Gln Leu Leu Glu Leu Ile Arg Ile Asn His Gly
305                 310                 315                 320

Leu Leu Val Ser Ile Gly Val Ser His Pro Gly Leu Glu Leu Ile Lys
```

325 330 335

Asn Leu Ser Asp Asp Leu Arg Ile Gly Ser Thr Lys Leu Thr Gly Ala
340 345 350

Gly Gly Gly Gly Cys Ser Leu Thr Leu Leu Arg Arg Asp Ile Thr Gln
355 360 365

Glu Gln Ile Asp Ser Phe Lys Lys Lys Leu Gln Asp Asp Phe Ser Tyr
370 375 380

Glu Thr Phe Glu Thr Asp Leu Gly Gly Thr Gly Cys Cys Leu Leu Ser
385 390 395 400

Ala Lys Asn Leu Asn Lys Asp Leu Lys Ile Lys Ser Leu Val Phe Gln
405 410 415

Leu Phe Glu Asn Lys Thr Thr Thr Lys Gln Gln Ile Asp Asp Leu Leu
420 425 430

Leu Pro Gly Asn Thr Asn Leu Pro Trp Thr Ser
435 440

<210> 7
<211> 404
<212> PRT
<213> Schizosaccharomyces pombe

<400> 7

Met Ser Lys Ser Leu Ile Val Ser Ser Pro Gly Lys Thr Ile Leu Phe
1 5 10 15

Gly Glu His Ala Val Val Tyr Gly Ala Thr Ala Leu Ala Ala Ala Val
20 25 30

Ser Leu Arg Ser Tyr Cys Lys Leu Gln Thr Thr Asn Asn Asn Glu Ile
35 40 45

Val Ile Val Met Ser Asp Ile Gly Thr Glu Arg Arg Trp Asn Leu Gln
50 55 60

Ser Leu Pro Trp Gln His Val Thr Val Glu Asn Val Gln His Pro Ala
65 70 75 80

Ser Ser Pro Asn Leu Asp Leu Leu Gln Gly Leu Gly Glu Leu Leu Lys
85 90 95

Asn Glu Glu Asn Gly Leu Ile His Ser Ala Met Leu Cys Thr Leu Tyr

```
                100                 105                 110

Leu Phe Thr Ser Leu Ser Ser Pro Ser Gln Gly Cys Thr Leu Thr Ile
        115                 120                 125

Ser Ser Gln Val Pro Leu Gly Ala Gly Leu Gly Ser Ser Ala Thr Ile
    130                 135                 140

Ser Val Val Val Ala Thr Ser Leu Leu Leu Ala Phe Gly Asn Ile Glu
145                 150                 155                 160

Pro Pro Ser Ser Asn Ser Leu Gln Asn Asn Lys Ala Leu Ala Leu Ile
                165                 170                 175

Glu Ala Trp Ser Phe Leu Gly Glu Cys Cys Ile His Gly Thr Pro Ser
            180                 185                 190

Gly Ile Asp Asn Ala Val Ala Thr Asn Gly Gly Leu Ile Ala Phe Arg
        195                 200                 205

Lys Ala Thr Ala His Gln Ser Ala Met Lys Glu Phe Leu Lys Pro Lys
    210                 215                 220

Asp Thr Leu Ser Val Met Ile Thr Asp Thr Lys Gln Pro Lys Ser Thr
225                 230                 235                 240

Lys Lys Leu Val Gln Gly Val Phe Glu Leu Lys Glu Arg Leu Pro Thr
                245                 250                 255

Val Ile Asp Ser Ile Ile Asp Ala Ile Asp Gly Ile Ser Lys Ser Ala
            260                 265                 270

Val Leu Ala Leu Thr Ser Glu Ser Asp Lys Asn Ser Ser Ala Lys Lys
        275                 280                 285

Leu Gly Glu Phe Ile Val Leu Asn Gln Lys Leu Leu Glu Cys Leu Gly
    290                 295                 300

Val Ser His Tyr Ser Ile Asp Arg Val Leu Gln Ala Thr Lys Ser Ile
305                 310                 315                 320

Gly Trp Thr Lys Leu Thr Gly Ala Gly Gly Gly Gly Cys Thr Ile Thr
                325                 330                 335

Leu Leu Thr Pro Glu Cys Lys Glu Glu Glu Phe Lys Leu Cys Lys Glu
            340                 345                 350
```

```
Ser Leu Leu Ala His Lys Asn Ser Ile Tyr Asp Val Gln Leu Gly Gly
        355                 360                 365

Pro Gly Val Ser Val Val Thr Asp Ser Asp Ser Phe Phe Pro Gln Tyr
    370                 375                 380

Glu Ser Asp Phe Asp Phe Lys Lys Leu Asn Leu Leu Ser Lys Phe Asn
385                 390                 395                 400

Lys Tyr Tyr Ile
```

<210> 8
<211> 335
<212> PRT
<213> Pyrococcus abyssi

<400> 8

```
Met Pro Arg Leu Val Leu Ala Ser Ala Pro Ala Lys Ile Ile Leu Phe
1               5                   10                  15

Gly Glu His Ser Val Val Tyr Gly Lys Pro Ala Ile Ala Ser Ala Ile
            20                  25                  30

Asp Leu Arg Thr Tyr Val Arg Ala Glu Phe Asn Asp Ser Gly Asn Ile
        35                  40                  45

Lys Ile Glu Ala His Asp Ile Lys Thr Pro Gly Leu Ile Val Ser Phe
    50                  55                  60

Ser Glu Asp Lys Ile Tyr Phe Glu Thr Asp Tyr Gly Lys Ala Ala Glu
65                  70                  75                  80

Val Leu Ser Tyr Val Arg His Ala Ile Glu Leu Val Leu Glu Glu Ala
                85                  90                  95

Asp Lys Arg Thr Gly Val Ser Val Ser Ile Thr Ser Gln Ile Pro Val
            100                 105                 110

Gly Ala Gly Leu Gly Ser Ser Ala Ala Val Ala Val Ala Thr Ile Gly
        115                 120                 125

Ala Val Ser Lys Leu Leu Asp Leu Glu Leu Ser Lys Glu Glu Ile Ala
    130                 135                 140

Lys Met Gly His Lys Val Glu Leu Leu Val Gln Gly Ala Ser Ser Gly
145                 150                 155                 160
```

```
Ile Asp Pro Thr Val Ser Ala Ile Gly Gly Phe Leu Tyr Tyr Lys Gln
                165                 170                 175

Gly Glu Phe Glu His Leu Pro Phe Val Glu Leu Pro Ile Val Val Gly
            180                 185                 190

Tyr Thr Gly Ser Ser Gly Ser Thr Lys Glu Leu Val Ala Met Val Arg
        195                 200                 205

Arg Arg Tyr Glu Glu Met Pro Glu Leu Ile Glu Pro Ile Leu Glu Ser
    210                 215                 220

Met Gly Lys Leu Val Asp Lys Ala Lys Glu Val Ile Ile Ser Lys Leu
225                 230                 235                 240

Asp Glu Glu Glu Lys Phe Leu Lys Leu Gly Glu Leu Met Asn Ile Asn
                245                 250                 255

His Gly Leu Leu Asp Ala Leu Gly Val Ser Thr Lys Lys Leu Ser Glu
            260                 265                 270

Leu Val Tyr Ala Ala Arg Thr Ala Gly Ala Ile Gly Ala Lys Leu Thr
        275                 280                 285

Gly Ala Gly Gly Gly Gly Cys Met Tyr Ala Leu Ala Pro Gly Lys Gln
    290                 295                 300

Arg Glu Val Ala Thr Ala Ile Lys Ile Ala Gly Gly Thr Pro Met Ile
305                 310                 315                 320

Thr Arg Ile Ser Lys Glu Gly Leu Arg Ile Glu Glu Val Arg Glu
                325                 330                 335
```

<210> 9
<211> 335
<212> PRT
<213> Pyrococcus horikoshii

<400> 9

```
Met Val Lys Tyr Val Leu Ala Ser Ala Pro Ala Lys Val Ile Leu Phe
1               5                   10                  15

Gly Glu His Ser Val Val Tyr Gly Lys Pro Ala Ile Ala Ser Ala Ile
            20                  25                  30

Glu Leu Arg Thr Tyr Val Arg Ala Gln Phe Asn Asp Ser Gly Asn Ile
        35                  40                  45
```

```
Lys Ile Glu Ala His Asp Ile Lys Thr Pro Gly Leu Ile Val Ser Phe
    50                  55                  60

Ser Glu Asp Lys Ile Tyr Phe Glu Thr Asp Tyr Gly Lys Ala Ala Glu
65                  70                  75                  80

Val Leu Ser Tyr Val Arg Tyr Ala Ile Glu Leu Ala Leu Glu Glu Ser
                85                  90                  95

Asp Lys Arg Val Gly Ile Asp Val Ser Ile Thr Ser Gln Ile Pro Val
            100                 105                 110

Gly Ala Gly Leu Gly Ser Ser Ala Ala Val Ala Val Ala Thr Ile Gly
        115                 120                 125

Ala Val Ser Arg Leu Leu Gly Leu Glu Leu Ser Lys Glu Glu Ile Ala
    130                 135                 140

Lys Leu Gly His Lys Val Glu Leu Leu Val Gln Gly Ala Ser Ser Gly
145                 150                 155                 160

Ile Asp Pro Thr Val Ser Ala Val Gly Gly Phe Leu Tyr Tyr Lys Gln
                165                 170                 175

Gly Lys Phe Glu Pro Leu Pro Phe Met Glu Leu Pro Ile Val Val Gly
            180                 185                 190

Tyr Thr Gly Ser Thr Gly Ser Thr Lys Glu Leu Val Ala Met Val Arg
        195                 200                 205

Lys Arg Tyr Glu Glu Met Pro Glu Leu Val Glu Pro Ile Leu Glu Ala
    210                 215                 220

Met Gly Lys Leu Val Asp Lys Ala Lys Glu Ile Ile Leu Ser Lys Leu
225                 230                 235                 240

Asp Glu Glu Glu Lys Leu Thr Lys Leu Gly Glu Leu Met Asn Ile Asn
                245                 250                 255

His Gly Leu Leu Asp Ala Leu Gly Val Ser Thr Lys Lys Leu Gly Glu
            260                 265                 270

Leu Val Tyr Ala Ala Arg Thr Ala Gly Ala Ile Gly Ala Lys Leu Thr
        275                 280                 285

Gly Ala Gly Gly Gly Gly Cys Met Tyr Ala Leu Ala Pro Gly Arg Gln
    290                 295                 300
```

```
Arg Glu Val Ala Thr Ala Ile Lys Ile Ala Gly Gly Ile Pro Met Ile
305                 310                 315                 320

Thr Arg Val Ser Arg Glu Gly Leu Arg Ile Glu Glu Val Ser Arg
                325                 330                 335
```

<210> 10
<211> 334
<212> PRT
<213> Pyrococcus furiosus

<400> 10

```
Met Lys Val Ile Ala Ser Ala Pro Ala Lys Val Ile Leu Phe Gly Glu
1               5                   10                  15

His Ser Val Val Tyr Gly Lys Pro Ala Ile Ala Ala Ala Ile Asp Leu
            20                  25                  30

Arg Thr Phe Val Glu Ala Glu Leu Ile Arg Glu Lys Lys Ile Arg Ile
        35                  40                  45

Glu Ala His Asp Ile Lys Val Pro Gly Leu Thr Val Ser Phe Ser Glu
    50                  55                  60

Asn Glu Ile Tyr Phe Glu Thr Asp Tyr Gly Lys Ala Ala Glu Val Leu
65                  70                  75                  80

Ser Tyr Val Arg Glu Ala Ile Asn Leu Val Leu Glu Glu Ala Asp Lys
                85                  90                  95

Lys Asn Val Gly Ile Lys Val Ser Ile Thr Ser Gln Ile Pro Val Gly
            100                 105                 110

Ala Gly Leu Gly Ser Ser Ala Ala Val Ala Val Ala Thr Ile Gly Ala
        115                 120                 125

Val Ser Lys Leu Leu Gly Leu Glu Leu Ser Lys Glu Glu Ile Ala Lys
    130                 135                 140

Met Gly His Lys Thr Glu Leu Leu Val Gln Gly Ala Ser Ser Gly Ile
145                 150                 155                 160

Asp Pro Thr Val Ser Ala Ile Gly Gly Phe Ile Phe Tyr Glu Lys Gly
                165                 170                 175

Lys Phe Glu His Leu Pro Phe Met Glu Leu Pro Ile Val Val Gly Tyr
            180                 185                 190
```

```
Thr Gly Ser Ser Gly Pro Thr Lys Glu Leu Val Ala Met Val Arg Lys
        195                 200                 205

Arg Tyr Glu Glu Met Pro Glu Leu Ile Val Pro Ile Leu Glu Ala Met
    210                 215                 220

Gly Lys Val Val Glu Lys Ala Lys Asp Val Ile Leu Ser Asn Val Asp
225                 230                 235                 240

Lys Glu Glu Lys Phe Glu Arg Leu Gly Val Leu Met Asn Ile Asn His
                245                 250                 255

Gly Leu Leu Asp Ala Leu Gly Val Ser Thr Lys Lys Leu Ser Glu Leu
            260                 265                 270

Val Tyr Ala Ala Arg Val Ala Gly Ala Leu Gly Ala Lys Ile Thr Gly
        275                 280                 285

Ala Gly Gly Gly Gly Cys Met Tyr Ala Leu Ala Pro Asn Lys Gln Arg
    290                 295                 300

Glu Val Ala Thr Ala Ile Arg Ile Ala Gly Gly Thr Pro Met Ile Thr
305                 310                 315                 320

Glu Ile Ser Arg Glu Gly Leu Lys Ile Glu Glu Val Ile Lys
                325                 330
```

<210> 11
<211> 303
<212> PRT
<213> Methanobacterium thermoautotrophicum

<400> 11

```
Met Lys Ser Ser Ala Ser Ala Pro Ala Lys Ala Ile Leu Phe Gly Glu
1               5                   10                  15

His Ala Val Val Tyr Ser Lys Pro Ala Ile Ala Ala Ala Ile Asp Arg
            20                  25                  30

Arg Val Thr Val Thr Val Ser Glu Ser Ser Ser Thr His Val Thr Ile
        35                  40                  45

Pro Ser Leu Gly Ile Arg His Ser Ser Glu Arg Pro Ser Gly Gly Ile
    50                  55                  60

Leu Asp Tyr Ile Gly Arg Cys Leu Glu Leu Tyr His Asp Ala Ser Pro
65                  70                  75                  80
```

```
Leu Asp Ile Arg Val Glu Met Glu Ile Pro Ala Gly Ser Gly Leu Gly
                85                  90                  95

Ser Ser Ala Ala Leu Thr Val Ala Leu Ile Gly Ala Leu Asp Arg Tyr
            100                 105                 110

His Gly Arg Asp His Gly Pro Gly Glu Thr Ala Ala Arg Ala His Arg
        115                 120                 125

Val Glu Val Asp Val Gln Gly Ala Ala Ser Pro Leu Asp Thr Ala Ile
    130                 135                 140

Ser Thr Tyr Gly Gly Leu Val Tyr Leu Asp Ser Gln Arg Arg Val Arg
145                 150                 155                 160

Gln Phe Glu Ala Asp Leu Gly Asp Leu Val Ile Ala His Leu Asp Tyr
                165                 170                 175

Ser Gly Glu Thr Ala Arg Met Val Ala Gly Val Ala Glu Arg Phe Arg
            180                 185                 190

Arg Phe Pro Asp Ile Met Gly Arg Ile Met Asp Thr Val Glu Ser Ile
        195                 200                 205

Thr Asn Thr Ala Tyr Arg Glu Leu Leu Arg Asn Asn Thr Glu Pro Leu
    210                 215                 220

Gly Glu Leu Met Asn Leu Asn Gln Gly Leu Leu Asp Ser Met Gly Val
225                 230                 235                 240

Ser Thr Arg Glu Leu Ser Met Met Val Tyr Glu Ala Arg Asn Ala Gly
                245                 250                 255

Ala Ala Gly Ser Lys Ile Thr Gly Ala Gly Gly Gly Gly Ser Ile Ile
            260                 265                 270

Ala His Cys Pro Gly Cys Val Asp Asp Val Val Thr Ala Leu Asn Arg
        275                 280                 285

Asn Trp Lys Ala Met Arg Ala Glu Phe Ser Val Lys Gly Leu Ile
    290                 295                 300
```

<210> 12
<211> 284
<212> PRT
<213> Archaeoglobus fulgidus

<400> 12

```
Met Ile Ala Ser Ala Pro Gly Lys Ile Ile Leu Phe Gly Glu His Ala
1               5                  10                  15

Val Val Tyr Gly Arg His Ala Val Val Ser Ala Ile Asn Leu Arg Cys
            20                  25                  30

Arg Val Ser Val Arg Lys Ser Asp Arg Phe Leu Ile Arg Ser Ser Leu
        35                  40                  45

Gly Glu Ser Gly Leu Asp Tyr Gln Arg His Pro Tyr Val Val Gln Ala
    50                  55                  60

Val Lys Arg Phe Gly Glu Leu Arg Asn Ile Pro Gly Ala Glu Ile Glu
65                  70                  75                  80

Ile Glu Ser Glu Ile Pro Ile Gly Ser Gly Leu Gly Ser Ser Ala Ala
                85                  90                  95

Val Ile Val Ala Thr Ile Ala Ala Leu Asn Ala Glu Phe Asp Gly Asp
            100                 105                 110

Met Asp Lys Glu Ala Ile Phe Gln Met Ala Lys Gln Val Glu Ile Asp
        115                 120                 125

Val Gln Gly Arg Ala Ser Gly Ile Asp Pro Phe Ile Ser Thr Phe Gly
    130                 135                 140

Gly Ser Trp Leu Phe Pro Glu Arg Arg Lys Val Glu Met Pro Phe Lys
145                 150                 155                 160

Phe Phe Val Ile Asn Phe Gly Ser Arg Ser Thr Ala Glu Met Val Ala
                165                 170                 175

Lys Val Ala Glu Leu Arg Glu Arg His Pro Glu Val Val Asp Lys Ile
            180                 185                 190

Phe Asp Ala Ile Asp Ala Ile Ser Leu Glu Ala Ser Asp Val Gly Ser
        195                 200                 205

Ala Glu Arg Leu Glu Glu Leu Ile Ala Ile Asn Gln Ser Leu Leu Arg
    210                 215                 220

Ala Ile Gly Val Ser Asn Pro Glu Ile Asp Arg Thr Ile Ala Glu Leu
225                 230                 235                 240

Glu Arg Met Gly Leu Asn Ala Lys Ile Thr Gly Ala Gly Gly Gly Gly
                245                 250                 255
```

```
Cys Ile Phe Gly Leu Phe Lys Gly Glu Lys Pro Lys Gly Ser Phe Ile
            260                 265                 270

Val Glu Pro Glu Lys Glu Gly Val Arg Ile Glu Glu
        275                 280
```

<210> 13
<211> 312
<212> PRT
<213> Methanococcus jannaschii

<400> 13

```
Met Ile Ile Glu Thr Pro Ser Lys Val Ile Leu Phe Gly Glu His Ala
1               5                   10                  15

Val Val Tyr Gly Tyr Arg Ala Ile Ser Met Ala Ile Asp Leu Thr Ser
            20                  25                  30

Thr Ile Glu Ile Lys Glu Thr Gln Glu Asp Glu Ile Ile Leu Asn Leu
        35                  40                  45

Asn Asp Leu Asn Lys Ser Leu Gly Leu Asn Leu Asn Glu Ile Lys Asn
    50                  55                  60

Ile Asn Pro Asn Asn Phe Gly Asp Phe Lys Tyr Cys Leu Cys Ala Ile
65                  70                  75                  80

Lys Asn Thr Leu Asp Tyr Leu Asn Ile Glu Pro Lys Thr Gly Phe Lys
                85                  90                  95

Ile Asn Ile Ser Ser Lys Ile Pro Ile Ser Cys Gly Leu Gly Ser Ser
            100                 105                 110

Ala Ser Ile Thr Ile Gly Thr Ile Lys Ala Val Ser Gly Phe Tyr Asn
        115                 120                 125

Lys Glu Leu Lys Asp Asp Glu Ile Ala Lys Leu Gly Tyr Met Val Glu
    130                 135                 140

Lys Glu Ile Gln Gly Lys Ala Ser Ile Thr Asp Thr Ser Thr Ile Thr
145                 150                 155                 160

Tyr Lys Gly Ile Leu Glu Ile Lys Asn Asn Lys Phe Arg Lys Ile Lys
                165                 170                 175

Gly Glu Phe Glu Glu Phe Leu Lys Asn Cys Lys Phe Leu Ile Val Tyr
            180                 185                 190
```

```
Ala Glu Lys Arg Lys Lys Lys Thr Ala Glu Leu Val Asn Glu Val Ala
        195                 200                 205

Lys Ile Glu Asn Lys Asp Glu Ile Phe Lys Glu Ile Asp Lys Val Ile
    210                 215                 220

Asp Glu Ala Leu Lys Ile Lys Asn Lys Glu Asp Phe Gly Lys Leu Met
225                 230                 235                 240

Thr Lys Asn His Glu Leu Leu Lys Lys Leu Asn Ile Ser Thr Pro Lys
                245                 250                 255

Leu Asp Arg Ile Val Asp Ile Gly Asn Arg Phe Gly Phe Gly Ala Lys
            260                 265                 270

Leu Thr Gly Ala Gly Gly Gly Gly Cys Val Ile Ile Leu Val Asn Glu
        275                 280                 285

Glu Lys Glu Lys Glu Leu Leu Lys Glu Leu Asn Lys Glu Asp Val Arg
    290                 295                 300

Ile Phe Asn Cys Arg Met Met Asn
305                 310
```

<210> 14
<211> 324
<212> PRT
<213> Aeropyrum pernix

<400> 14

```
Met Arg Arg Ala Ala Arg Ala Ser Ala Pro Gly Lys Val Ile Ile Val
1               5                   10                  15

Gly Glu His Phe Val Val Arg Gly Ser Leu Ala Ile Val Ala Ala Ile
            20                  25                  30

Gly Arg Arg Leu Arg Val Thr Val Arg Ser Gly Gly Lys Gly Ile Val
        35                  40                  45

Leu Glu Ser Ser Met Leu Gly Arg His Ser Ala Pro Leu Pro Gly Gln
    50                  55                  60

Gly Ala Ala Ala Lys Val Ser Pro Val Leu Glu Pro Tyr Ile Ala Val
65                  70                  75                  80

Leu Arg Ser Leu Ala Ala Arg Gly Tyr Ser Val Val Pro His Thr Ile
                85                  90                  95
```

Leu Val Glu Ser Gly Ile Pro Pro Arg Ala Gly Leu Gly Ser Ser Ala
100 105 110

Ala Ser Met Val Ala Tyr Ala Leu Ser Tyr Ser Ala Met His Gly Asp
115 120 125

Pro Leu Ser Ala Glu Asp Leu Tyr Ser Val Ala Met Glu Gly Glu Lys
130 135 140

Ile Ala His Gly Lys Pro Ser Gly Val Asp Val Thr Ile Ala Val Arg
145 150 155 160

Gly Gly Val Leu Ala Tyr Arg Arg Gly Glu Asn Pro Val Asp Ile Arg
165 170 175

Pro Gly Leu Thr Gly Val Thr Leu Leu Val Ala Asp Thr Gly Val Glu
180 185 190

Arg Arg Thr Arg Asp Val Val Glu His Val Leu Ser Ile Ala Asp Ala
195 200 205

Leu Gly Glu Ala Ser Thr Tyr Ile Tyr Arg Ala Ala Asp Leu Ile Ala
210 215 220

Arg Glu Ala Leu His Ala Ile Glu Lys Gly Asp Ala Glu Arg Leu Gly
225 230 235 240

Leu Ile Met Asn Ala Ala Gln Gly Leu Leu Ser Ser Leu Gly Ala Ser
245 250 255

Ser Leu Glu Ile Glu Thr Leu Val Tyr Arg Met Arg Ser Ala Gly Ala
260 265 270

Leu Gly Ala Lys Leu Thr Gly Ala Gly Trp Gly Gly Cys Val Ile Gly
275 280 285

Leu Phe Lys Glu Gly Glu Val Glu Arg Gly Leu Glu Ser Val Val Glu
290 295 300

Ser Ser Ser Gln Ala Phe Thr Ala Ser Ile Ala Glu Glu Gly Ala Arg
305 310 315 320

Leu Glu Glu Phe

<210> 15

<211> 387
<212> PRT
<213> Artificial Sequence

<220>
<223> His6-Paracoccus zeaxanthinifaciens

<400> 15

```
Met Arg Gly Ser His His His His His His Ser Thr Gly Arg Pro Glu
1               5                  10                  15

Ala Gly Ala His Ala Pro Gly Lys Leu Ile Leu Ser Gly Glu His Ser
            20                  25                  30

Val Leu Tyr Gly Ala Pro Ala Leu Ala Met Ala Ile Ala Arg Tyr Thr
        35                  40                  45

Glu Val Trp Phe Thr Pro Leu Gly Ile Gly Glu Gly Ile Arg Thr Thr
    50                  55                  60

Phe Ala Asn Leu Ser Gly Gly Ala Thr Tyr Ser Leu Lys Leu Leu Ser
65                  70                  75                  80

Gly Phe Lys Ser Arg Leu Asp Arg Arg Phe Glu Gln Phe Leu Asn Gly
                85                  90                  95

Asp Leu Lys Val His Lys Val Leu Thr His Pro Asp Asp Leu Ala Val
            100                 105                 110

Tyr Ala Leu Ala Ser Leu Leu His Asp Lys Pro Pro Gly Thr Ala Ala
        115                 120                 125

Met Pro Gly Ile Gly Ala Met His His Leu Pro Arg Pro Gly Glu Leu
    130                 135                 140

Gly Ser Arg Thr Glu Leu Pro Ile Gly Ala Gly Met Gly Ser Ser Ala
145                 150                 155                 160

Ala Ile Val Ala Ala Thr Thr Val Leu Phe Glu Thr Leu Leu Asp Arg
                165                 170                 175

Pro Lys Thr Pro Glu Gln Arg Phe Asp Arg Val Arg Phe Cys Glu Arg
            180                 185                 190

Leu Lys His Gly Lys Ala Gly Pro Ile Asp Ala Ala Ser Val Val Arg
        195                 200                 205

Gly Gly Leu Val Arg Val Gly Gly Asn Gly Pro Gly Ser Ile Ser Ser
    210                 215                 220
```

```
Phe Asp Leu Pro Glu Asp His Asp Leu Val Ala Gly Arg Gly Trp Tyr
225                 230                 235                 240

Trp Val Leu His Gly Arg Pro Val Ser Gly Thr Gly Glu Cys Val Ser
                245                 250                 255

Ala Val Ala Ala Ala His Gly Arg Asp Ala Ala Leu Trp Asp Ala Phe
            260                 265                 270

Ala Val Cys Thr Arg Ala Leu Glu Ala Ala Leu Leu Ser Gly Gly Ser
        275                 280                 285

Pro Asp Ala Ala Ile Thr Glu Asn Gln Arg Leu Leu Glu Arg Ile Gly
    290                 295                 300

Val Val Pro Ala Ala Thr Gln Ala Leu Val Ala Gln Ile Glu Glu Ala
305                 310                 315                 320

Gly Gly Ala Ala Lys Ile Cys Gly Ala Gly Ser Val Arg Gly Asp His
                325                 330                 335

Gly Gly Ala Val Leu Val Arg Ile Asp Asp Ala Gln Ala Met Ala Ser
            340                 345                 350

Val Met Ala Arg His Pro Asp Leu Asp Trp Ala Pro Leu Arg Met Ser
        355                 360                 365

Arg Thr Gly Ala Ala Pro Gly Pro Ala Pro Arg Ala Gln Pro Leu Pro
    370                 375                 380

Gly Gln Gly
385
```

<210> 16
<211> 1137
<212> DNA
<213> Paracoccus zeaxanthinifaciens

<400> 16

```
atgtcgaccg gcaggcctga agcaggcgcc catgccccgg gcaagctgat cctgtccggg     60

gaacattccg tgctctatgg tgcgcccgcg cttgccatgg ccatcgcccg ctataccgag    120

gtgtggttca cgccgcttgg cattggcgag gggatacgca cgacattcgc caatctctcg    180

ggcggggcga cctattcgct gaagctgctg tcggggttca agtcgcggct ggaccgccgg    240

ttcgagcagt tcctgaacgg cgacctaaag gtgcacaagg tcctgaccca tcccgacgat    300

ctggcggtct atgcgctggc gtcgcttctg cacgacaagc cgccggggac cgccgcgatg    360
```

ccgggcatcg gcgcgatgca ccacctgccg cgaccgggtg agctgggcag ccggacggag 420

ctgcccatcg gcgcgggcat ggggtcgtct gcggccatcg tcgcggccac cacggtcctg 480

ttcgagacgc tgctggaccg gcccaagacg cccgaacagc gcttcgaccg cgtccgcttc 540

tgcgagcggt tgaagcacgg caaggccggt cccatcgacg cggccagcgt cgtgcgcggc 600

gggcttgtcc gcgtgggcgg gaacgggccg ggttcgatca gcagcttcga tttgcccgag 660

gatcacgacc ttgtcgcggg acgcggctgg tactgggtac tgcacgggcg ccccgtcagc 720

gggaccggcg aatgcgtcag cgcggtcgcg gcggcgcatg gtcgcgatgc ggcgctgtgg 780

gacgccttcg cagtctgcac ccgcgcgttg gaggccgcgc tgctgtctgg gggcagcccc 840

gacgccgcca tcaccgagaa ccagcgcctg ctggaacgca tcggcgtcgt gccggcagcg 900

acgcaggccc tcgtggccca gatcgaggag gcgggtggcg cggccaagat ctgcggcgca 960

ggttccgtgc ggggcgatca cggcggggcg gtcctcgtgc ggattgacga cgcgcaggcg 1020

atggcttcgg tcatggcgcg ccatcccgac ctcgactggg cgcccctgcg catgtcgcgc 1080

acgggggcgg cacccggccc cgcgccgcgt gcgcaaccgc tgccggggca gggctga 1137

<210> 17
<211> 1191
<212> DNA
<213> homo sapiens

<400> 17
atgttgtcag aagtcctact ggtgtctgct ccggggaaag tcatccttca tggagaacat 60

gccgtggtac atggcaaggt agcactggct gtatccttga acttgagaac attcctccgg 120

cttcaacccc acagcaatgg gaaagtggac ctcagcttac ccaacattgg tatcaagcgg 180

gcctgggatg tggccaggct tcagtcactg gacacaagct ttctggagca aggtgatgtc 240

acaacaccca cctcagagca agtggagaag ctaaaggagg ttgcaggctt gcctgacgac 300

tgtgctgtca ccgagcgcct ggctgtgctg gcctttcttt acttatacct gtccatctgc 360

cggaagcaga gggccctgcc gagcctggat atcgtagtgt ggtcggagct gccccccggg 420

gcgggcttgg gctccagcgc cgcctactcg gtgtgtctgg cagcagccct cctgactgtg 480

tgcgaggaga tcccaaaccc gctgaaggac ggggattgcg tcaacaggtg gaccaaggag 540

gatttggagc taattaacaa gtgggccttc caaggggaga gaatgattca cgggaacccc 600

tccggagtgg acaatgctgt cagcacctgg ggaggagccc tccgatacca tcaagggaag 660

atttcatcct taaagaggtc gccagctctc cagatcctgc tgaccaacac caaagtccct 720

cgcaatacca gggcccttgt ggctggcgtc agaaacaggc tgctcaagtt cccagagatc 780

gtggcccccc tcctgacctc aatagatgcc atctccctgg agtgtgagcg cgtgctggga 840

gagatggggg aagccccagc cccggagcag tacctcgtgc tggaagagct cattgacatg 900

aaccagcacc atctgaatgc cctcggcgtg ggccacgcct ctctggacca gctctgccag 960

gtgaccaggg cccgcggact tcacagcaag ctgactggcg caggcggtgg tggctgtggc 1020

atcacactcc tcaagccagg gctggagcag ccagaagtgg aggccacgaa gcaggccctg 1080

accagctgtg gctttgactg cttggaaacc agcatcggtg cccccggcgt ctccatccac 1140

tcagccacct ccctggacag ccgagtccag caagccctgg atggcctctg a 1191

<210> 18
<211> 1188
<212> DNA
<213> mouse

<400> 18
atgttgtcag aagccctgct ggtgtccgcc ccggggaagg tcatcctcca tggagaacac 60

gctgtggtcc atggcaaggt cgctctggca gcggccttga acttgagaac tttcctcctg 120

ctgcgaccgc agagcaatgg gaaagtgagc gtcaatttac ccaacatcgg tattaagcag 180

gtgtgggatg tgggcatgct tcagcgactg gacacgagct ttcttgagca aggtgatgtc 240

tcggtaccca ccttggagca actggagaag ctaaagaaga tgggggacct ccccagagac 300

cgtgcaggca atgaaggcat ggctctgctt gcctttctct acctgtacct ggcaatctgc 360

cggaagcaga ggacactccc gagcctggac atggtggtgt ggtcggaact tccccccggg 420

gcaggcttgg gctccagcgc cgcctactct gtgtgtctgg cagccgccct cctgactgcc 480

tgtgaggagg tctccaaccc gctcaaggac ggggtctccg tcagcaggtg gcccgaggaa 540

gatctgaagt caatcaacaa gtgggccttc gaaggggaga gagtgatcca tgggaaccct 600

tctggtgtgg acaatgccgt cagcacctgg ggcggagccc tgcgcttcca gcaagggacg 660

atgtcttcct tgaagagcct cccgtctctg cagatcctgc tcaccaacac caaggtcccg 720

cggagtacca aggcccttgt ggctgctgtc agaagcaggc tgaccaagtt ccctgagatt 780

gtggccccgc tgctgacctc cattgacgca atatccctgg agtgtgagcg cgtgctaggg 840

gagatggtgg cagctccagt tccggaacag tacctcgtac tagaagagct gatagacatg 900

aaccagcacc atctgaatgc tctcggggtg ggccacaact ccctggacca gctctgccaa 960

gtaacggcag cacacggact gcacagcaag ctgacgggcg ctggcggtgg cggctgtggc 1020

atcaccctcc tgaagccagg tctagagcaa gccacagtgg aggcagccaa gcaggccctg 1080

accagctgcg ggtttgactg ctgggagacc agcatcggcg cacccggagt ttccacacac 1140

tcagctgcag ctgtagggga ccctgtccga caagccctgg gcctctga 1188

<210> 19
<211> 1188
<212> DNA
<213> rat

<400> 19
atgttgtcag aagtcctgct ggtgtctgct ccagggaaag tcattctcca tggagaacat 60

gctgtggtcc atggcaaggt agctctggcg gtggccttga acttgagaac atttctcgtg 120

ctgcgaccgc agagcaatgg gaaagtgagc ctcaatttac caaacgtcgg tattaagcag 180

gtctgggatg tggccacact tcagctgctg gacacaggct ttcttgagca aggcgatgtc 240

ccggcaccca ccttggagca actggagaag ctgaagaagg tggcgggcct cccccgagac 300

tgtgtaggca acgaaggcct gtctctgctt gcctttctgt acctgtacct ggctatctgc 360

cggaaacaga ggacactccc aagcctggac atcatggtgt ggtcggaact gccccctggg 420

gcgggcttgg gctccagtgc agcctactcg gtgtgtgtgg cagccgccct cctgactgcc 480

tgtgaggagg tcaccaaccc gctcaaggac aggggctcca ttggcagttg gcccgaggag 540

gacctgaagt caattaacaa gtgggcctac gaggggggaga gagtgatcca tgggaacccc 600

tctggcgtgg acaattccgt cagcacctgg ggaggagccc tgcgctacca gcaagggaag 660

atgtcatcct tgaagaggct cccagctctg cagatcctgc tcaccaacac caaggtccca 720

cgaagcacca aggccctcgt ggctggcgtc agaagcaggc taatcaagtt ccctgagatc 780

atggccccgc tcctgacatc aattgacgca atctccctgg agtgtgagcg cgtgctggga 840

gagatggcgg ccgcaccagt cccagaacag taccttgtcc tagaagagct aatggacatg 900

aaccagcacc atctgaatgc ccttggtgtg ggccacgcct ccctggacca gctctgtcag 960

gtaacagcag cacatggact gcacagcaag ctgactggcg caggcggcgg cggctgtggc 1020

atcaccctcc tgaagccagg tctagagcga gcaaaagtgg aggccgccaa gcaggccctg 1080

accggctgcg ggtttgactg ctgggagacc agcattggag cgcctggggt ctccatgcac 1140

tcagccacct ccatagagga ccctgtccga caagccctgg gcctctga 1188

<210> 20
<211> 1137
<212> DNA
<213> Arabidopsis thaliana

<400> 20
atggaagtga aagctagagc tcctgggaag atcatacttg caggggaaca cgctgttgtt 60

catggatcca ccgctgtagc tgccgccatt gatctctaca cttacgttac tctccgcttt 120

cctcttccat cagctgagaa caatgatagg cttacacttc agctcaagga catttccttg 180

gagttttcat ggtccttagc cagaatcaaa gaagcgattc cttatgattc aagcactctc 240

tgccgttcta cgccggcttc atgttcagag gagaccctta aatcaattgc agttttggtt 300

gaagagcaaa atcttccaaa ggaaaagatg tggctctcct ctgggatctc cacgtttctc 360

tggttataca ccagaattat agggttcaat ccggctacag tagtcattaa ctctgagctt 420

ccatacgggt ctggcctcgg ttcatcagca gctttatgtg tagctctcac agctgctctc 480

cttgcttctt ctatttcaga gaaaacccgt ggtaacggtt ggtcatctct cgatgaaacc 540

aatcttgagt tgctaaataa atgggctttc gaaggcgaaa agatcatcca tgggaaacct 600

tctgggatag acaacaccgt cagtgcatac ggcaacatga tcaagttctg ctcaggcgag 660

ataactcggt tacaatccaa catgcctctg agaatgctaa ttaccaacac tagagttggg 720

cgaaacacaa aagctctggt ctctggtgtg tcacagagag cggtaagaca tcctgatgcg 780

atgaagtcag tgttcaacgc cgtggattct ataagcaaag agctcgctgc gatcattcag 840

tctaaagacg agacctcagt tacagaaaaa gaagagagaa taaaagaact catggagatg 900

aaccaaggtc tgctcctgtc aatgggggtt agccacagct caatcgaggc tgtgattcta 960

accacggtca agcacaagct tgtctccaaa cttacaggag ctggtggcgg cggctgcgtc 1020

ctcactctat taccaaccgg gacggtggtg gacaaagtgg tggaggagct cgagtccagc 1080

ggttttcagt gtttcacggc attgattggt ggtaacggag ctcagatttg ctattga 1137

<210> 21
<211> 1332
<212> DNA
<213> Saccharomyces cerevisiae

<400> 21
atgtcattac cgttcttaac ttctgcaccg ggaaaggtta ttatttttgg tgaacactct 60

gctgtgtaca acaagcctgc cgtcgctgct agtgtgtctg cgttgagaac ctacctgcta 120

ataagcgagt catctgcacc agatactatt gaattggact tcccggacat tagctttaat 180

cataagtggt ccatcaatga tttcaatgcc atcaccgagg atcaagtaaa ctcccaaaaa 240

ttggccaagg ctcaacaagc caccgatggc ttgtctcagg aactcgttag tcttttggat 300

ccgttgttag ctcaactatc cgaatccttc cactaccatg cagcgttttg tttcctgtat 360

atgtttgttt gcctatgccc ccatgccaag aatattaagt tttctttaaa gtctacttta 420

cccatcggtg ctgggttggg ctcaagcgcc tctatttctg tatcactggc cttagctatg 480

gcctacttgg gggggttaat aggatctaat gacttggaaa agctgtcaga aaacgataag 540

catatagtga atcaatgggc cttcataggt gaaaagtgta ttcacggtac cccttcagga 600

atagataacg ctgtggccac ttatggtaat gccctgctat ttgaaaaaga ctcacataat 660

ggaacaataa acacaaacaa ttttaagttc ttagatgatt tcccagccat tccaatgatc 720

ctaacctata ctagaattcc aaggtctaca aaagatcttg ttgctcgcgt tcgtgtgttg 780

gtcaccgaga aatttcctga agttatgaag ccaattctag atgccatggg tgaatgtgcc 840

ctacaaggct tagagatcat gactaagtta agtaaatgta aaggcaccga tgacgaggct 900

gtagaaacta ataatgaact gtatgaacaa ctattggaat tgataagaat aaatcatgga 960

ctgcttgtct caatcggtgt ttctcatcct ggattagaac ttattaaaaa tctgagcgat 1020

gatttgagaa ttggctccac aaaacttacc ggtgctggtg gcggcggttg ctctttgact 1080

ttgttacgaa gagacattac tcaagagcaa attgacagct tcaaaaagaa attgcaagat 1140

gattttagtt acgagacatt tgaaacagac ttgggtggga ctggctgctg tttgttaagc 1200

gcaaaaaatt tgaataaaga tcttaaaatc aaatccctag tattccaatt atttgaaaat 1260

aaaactacca caaagcaaca aattgacgat ctattattgc caggaaacac gaatttacca 1320

tggacttcat aa 1332

<210> 22
<211> 1215
<212> DNA
<213> Schizosaccharomyces pombe

<400> 22
atgtcaaaat cgcttattgt ttcgtcgcca ggaaaaacga ttttgtttgg ggaacatgcc 60

gttgtatatg gagctactgc gttagctgca gctgtatcgt tacggagtta ctgtaaatta 120

cagacgacta ataacaatga aatagtaatt gtgatgagtg atatagggac cgaacgccga 180

tggaatcttc aatcgctacc ttggcagcat gtaacagtgg aaaacgttca gcacccggca 240

tcatctccca atctggacct tttacaagga ttaggagagc tattaaaaaa tgaagaaaac 300

ggacttattc actcagcaat gctttgtacc ctttacttgt tcacgtcttt gtcttctcct 360

tctcagggtt gtactttaac tattagctcc caagtacctt tgggtgctgg attaggtagt 420

agtgctacta tatcagttgt tgtcgctaca agtttactac tagcttttgg taatattgaa 480

cctcctagct caaattctct tcaaaacaac aaagcacttg cgttgataga ggcttggtct 540

tttctaggtg aatgctgtat tcatggaaca ccaagtggta ttgataatgc agtagcaaca 600

aatggaggac ttatcgcttt tcgtaaagct acagctcatc agagtgccat gaaagaattc 660

ttaaagccta aagatacctt atctgttatg atcactgata ccaaacaacc aaaaagtact 720

aaaaaacttg tacaaggagt ttttgaactg aaggaaagac taccaactgt gattgactca 780

ataatagatg caatcgatgg catatcaaag tctgccgtcc tcgcattgac ttcggagagc 840

gataaaaact cctccgctaa aaagttagga gagtttattg ttcttaatca aaaactctta 900

gaatgcttgg gtgtatccca ttattccatt gatcgcgttt tacaagccac taagtcaatt 960

ggatggacga agcttacagg tgccggtggt ggaggttgta cgattacttt attaacacct 1020

gagtgcaaag aagaggaatt taagttatgt aaagaatcac tattagccca taaaaattct 1080

atttacgatg ttcaattagg tggacctggt gtttcagtgg taaccgactc agattcattt 1140

ttccctcaat atgagtctga ctttgatttt aaaaaattga atttactcag caaatttaat 1200

aaatattata tttaa 1215

<210> 23
<211> 1008
<212> DNA
<213> Pyrococcus abyssi

<400> 23
atgccaaggt tagtgctggc gtcagctcca gcaaagataa tactcttcgg ggaacacagc 60

gttgtgtatg gaaagcctgc catagcatct gctattgact tgagaactta cgttagggcg 120

gagtttaatg attcgggaaa tataaagata gaagcccatg acataaaaac ccctgggcta 180

atagtttctt tttcagaaga caaaatttac ttcgagactg actatggaaa ggcagctgaa 240

gtgctgagtt acgttagaca cgccatagag ctcgtccttg aagaggctga taagaggact 300

ggggtcagcg tttcaataac gtctcaaatt ccagtaggtg ctggcctagg ttcttcagct 360

gccgtcgccg ttgctaccat cggtgccgtc tccaagttac ttgacctcga gcttagtaaa 420

gaggagatag ctaagatggg ccataaggtt gaactcctgg ttcagggagc ttcgagtggc 480

atagatccga cggtctcggc aataggaggg ttcttgtact ataagcaagg tgaatttgag 540

cacctaccat tcgtggagct tccaatagta gttggatata ccggctcaag tggctccaca 600

aaggaattag ttgcgatggt taggagaagg tacgaggaga tgcccgagtt aattgaaccc 660

attctagagt caatgggtaa gctcgtggat aaagctaagg aggtaataat atctaagctc 720

gatgaggagg aaaagttcct gaaattggga gagctcatga acataaatca tggccttctc 780

gatgccctag gtgtttcaac caaaaagcta agcgaactcg tctatgccgc tagaactgct 840

ggagcaattg gagccaagct aacgggggct gggggaggtg gatgcatgta cgctttagct 900

cctgggaagc agagggaggt tgctacggcc ataaagatag ctggcggaac tcccatgata 960

acgaggataa gcaaggaggg gcttagaata gaggaggtaa gggaatga 1008

<210> 24
<211> 1008
<212> DNA
<213> Pyrococcus horikoshii

<400> 24
tcatcttgaa acctcctcta ttctcaatcc ttccctactt accctcgtta tcattggaat 60

tccaccagct atctttattg ccgtggcgac ttccctctgt ctcccagggg ccaaggcata 120

catacaccct ccaccaccag cacccgtaag ttttgcccca attgcaccag ccgttctagc 180

agcgtaaact aactcaccaa gcttcttcgt agacaccccc aaagcatcta aaagcccatg 240

atttatgttc attaactctc caagcttagt aagcttctcc tcttcatcga gctttgaaag 300

tattatctcc ttggccttat ccactaattt tcccattgcc tccaatatag gctccacaag 360

ttcgggcatt tcctcgtacc ttttccttac cattgccact aattctttag ttgaaccagt 420

agagccggtg taaccaacga ctatgggaag ctccatgaat gggagaggct caaactttcc 480

ttgcttataa tagaggaagc ctcccactgc agaaactgta ggatcaatgc cacttgaagc 540

tccctgcact agaagttcaa ctttatgccc aagctttgct atttcctcct tactcaattc 600

aaggccaagt aacctagaga ccgcaccaat tgtagcaact gcaaccgctg ctgaggaacc 660

caatccggcc ccaactggaa tttgagaggt tattgaaacg tcaataccaa ccctcttatc 720

agactcctcc agggcaagtt ctatggcata cctcacatag ctcaaaactt cagcggcctt 780

tccatagtct gtttcaaagt atatcttatc ctcggaaaac gagactatca atccaggagt 840

ttttatatca tgagcttcta tttttatatt acccgaatca ttgaattgag ctctaacata 900

tgttctaagt tctatagcag aagctatagc cggttttcca tagactacgc tatgttcccc 960

gaagagtatt acctttgcag gagctgaagc taaaacgtac ttaaccat 1008

<210> 25
<211> 1005
<212> DNA
<213> Pyrococcus furiosus

<400> 25
tcatttaatt acctcctcaa tttttaaccc ttctctgctg atttcagtta tcataggagt 60

tccccctgca attcttatag ctgtagcaac ttctctctgc ttgttcggtg ccaaggcata 120

catacaacct cctcccccag ccccagttat cttagctcct agggctcccg caaccctagc 180

cgcgtacact agttcactca acttttttagt tgaaacaccc aaagcatcta aaagaccgtg 240

attaatgttc atcaaaaccc caagcctttc aaatttttcc tccttatcaa catttgaaag 300

tattacatcc ttggctttct cgacaacttt tcccatagcc tctaatatgg gaacaatcaa 360

ctcgggcatc tcctcatatc ttttcctaac cattgcaact aactccttag ttggacctga 420

ggagcccgtg tatccaacaa ctattggaag ctccatgaag ggaaggtgtt caaatttacc 480

cttttcatag aatataaacc cccctattgc agagaccgtt gggtctatac cacttgacgc 540

accttgaaca agtaattcag tcttgtgacc catcttagct atttcctcct tacttagctc 600

aagtcctagt aactttgata cagccccaat tgtggctact gccacggcag cagaagagcc 660

caatccagca ccaactggaa tttgagaagt tatgctaacc ttaataccaa catttttctt 720

atctgcctcc tctaaaacta aatttattgc ctctctaaca tagctcaaaa cttcagctgc 780

ttttccataa tctgtctcga aatatatctc attttcagaa aatgaaaccg taagtccagg 840

aactttaatg tcatgagctt caattcttat ttttttctcc cgaattagct cagcctccac 900

aaaagttcgt aaatcaatgg cagcagctat cgctggcttt ccgtaaacta cgctatgctc 960

tccaaaaaga ataacttttg cgggagctga ggctataact ttcat 1005

<210> 26
<211> 912

<212> DNA
<213> Methanobacterium thermoautotrophicum

<400> 26
ttgaagtcgt cggcatccgc acctgccaag gccattcttt ttggtgaaca cgcagtggtc 60

tacagcaagc cggcaatagc agccgccata gaccgcaggg tgactgtaac cgtaagtgaa 120

tccagcagca cccatgtaac catcccctcc ctgggtatac gccacagttc agagagacca 180

tccggtggca tcctggacta catcgggagg tgcctcgagc tttaccatga cgcatcaccc 240

cttgacatca gggtggagat ggagataccc gccggttcag gcctaggttc atcggctgca 300

ctcaccgttg cactgatagg tgccctcgac aggtaccatg gaagggatca tggacccggg 360

gagacagcag ccagggccca cagggtggag gttgatgtac agggagccgc cagccccctt 420

gacacagcca tcagcaccta tgggggcctt gtataccttg acagccagag gagggtgagg 480

cagtttgagg ccgacctggg ggaccttgta atagcacacc ttgactattc aggggaaaca 540

gccaggatgg ttgccggcgt agctgaaagg ttcaggagat tcccggatat catggggagg 600

ataatggaca cagttgagtc cataaccaat acagcataca gggaacttct aaggaacaac 660

acagaacccc tgggggagct catgaacctc aaccaggggc tgctggactc catgggcgtt 720

tccacacgtg aactttcaat gatggtctat gaggcaagga acgccggggc agcaggttca 780

aagatcacag gagccggcgg cggcgggagc ataatagccc actgcccggg atgtgtggat 840

gatgttgtca cggcccttaa caggaactgg aaagccatga gggcagagtt ttcggttaag 900

ggactcatct aa 912

<210> 27
<211> 855
<212> DNA
<213> Archaeoglobus fulgidus

<400> 27
tcactcttca attctgacac cctccttttc gggctcgact atgaagctcc ctttcggctt 60

ctctcccttg aacaggccga atatacaccc cccaccaccc gctccagtta ttttcgcatt 120

caaacccatc ctttcgagct ctgcaatggt tctgtcgatt tcgggattgc tcacccctat 180

cgccctcaga agcgactggt ttatggctat gagctcctca agtctttcag cgctgccaac 240

atcgctcgcc tcaagggaga ttgcgtcgat agcgtcaaat attttatcca caacctcagg 300

atgcctctct ctcagctcag caacttttgc gaccatctca gccgttgacc tgctgccaaa 360

gttgataacg aagaacttga acggcatctc gacctttcta cgctcaggga aaagccacga 420

acctccaaag gtggatatga acgggtctat tccgctcgcc cttccctgaa cgtcaatctc 480

aacctgcttc gccatctgga agatggcctc tttatccata tccccatcga actctgcatt 540

cagagcagct attgtggcaa caatcaccgc cgcagagctg cccagcccag agccaatcgg 600

tatttcgctc tcaatctcaa tttcagcacc cggaatattt ctaagctccc caaacctctt 660

aacagcctgc acgacgtagg gatgcctttg atagtccagc ccgctctctc ctaacgaaga 720

cctaatcaga aacctgtcag attttctcac cgacactctg catctaaggt taatcgctga 780

caccaccgca tgcctaccgt aaaccaccgc atgctcgccg aacagaatta tctttccggg 840

tgctgatgca atcat 855

<210> 28
<211> 939
<212> DNA
<213> Methanococcus jannaschii

<400> 28
atgataattg aaacaccatc aaaagttata ctattcggag agcatgcagt tgtttatggt 60

tatagagcta tatctatggc tattgattta acatcaacca tagaaataaa agaaacacaa 120

gaagatgaga taattttaaa cctaaatgac ttgaataaaa gcttaggttt gaacttaaat 180

gagataaaaa atatcaatcc aaataacttt ggagatttta aatactgcct ctgtgcaatt 240

aaaaacactt tagattattt aaatatagag ccaaaaactg gttttaaaat taacattagc 300

tcaaaaattc caataagttg tggtttggga agctctgcct caataacaat tggaactata 360

aaagctgtaa gtggatttta taataaagag cttaaagatg atgagattgc aaaacttgga 420

tatatggttg agaaagaaat ccaaggtaag gcaagcatta cagacacttc gacaataacg 480

tataaaggta tcttagaaat aaaaaacaac aagtttagaa aaattaaagg agagtttgaa 540

gaatttttaa aaaattgcaa gtttttaatt gtttatgctg aaaaaaggaa gaaaaaaact 600

gctgagttag ttaatgaagt tgccaagatt gaaaataaag atgagatatt taaagagata 660

gacaaagtta ttgatgaagc tttaaaaatc aaaaataaag aagattttgg gaaattgatg 720

actaaaaacc acgagttgtt aaaaaagcta aatatctcaa caccaaaact tgatagaatt 780

gtagatattg ggaatagatt tggttttggg gcaaaattaa ctggagctgg aggggggagga 840

tgtgtaataa tcttagttaa tgaagaaaaa gagaaagagc ttttaaaaga actaaataaa 900

gaagatgtaa ggatttttaa ctgcagaatg atgaattaa 939

<210> 29
<211> 975
<212> DNA
<213> Aeropyrum pernix

<400> 29
atgaggaggg ctgctagggc gtctgccccg gggaaagtta taatcgttgg agaacacttc 60

gtcgtcagag gctccctggc gatagtggcg gccataggca gaaggctccg cgtcaccgtg 120

agaagcgggg gcaaggggat tgtgcttgag agcagcatgc taggccgcca cagcgccccg 180

ctaccaggcc agggtgcagc ggctaaggta agccccgtcc tcgagccgta catagcagtg 240

ttgagaagtc tggctgcaag gggctatagc gtagtgcccc atacaatatt ggtggagagc 300

ggcatacccc ctagggcagg tctcggtagc agcgccgcca gcatggtagc ctatgctcta 360

tcatactcgg ccatgcatgg tgacccctc tcggctgagg acctctacag tgttgctatg 420

gagggcgaga agatagcgca tggtaagccg agcggtgttg acgtaaccat agccgttagg 480

gggggagtcc tggcttacag gaggggcgag aacccggtgg atataaggcc ggggcttaca 540

ggtgtcactc tgcttgttgc cgacacgggt gtcgagaggc gtactaggga tgttgtcgag 600

catgttctct ccattgcgga cgccttggga gaggcatcga cctacatata tagggcggca 660

gacttgatag cgagagaagc cctccatgcg atagaaaagg gagacgccga gaggctaggt 720

cttataatga atgcagccca gggccttctc tcatctcttg gggcgtcgtc actagaaata 780

gaaacactag tatatcggat gaggagtgcc ggggccctgg gtgcaaagct aacgggagct 840

ggatgggggg gctgtgtgat agggcttttc aaggagggtg aggtcgaacg ggggctagag 900

tctgtggtag agagttcaag ccaggctttc accgcgtcaa tagcagagga gggtgctaga 960

ctcgaggagt tctag 975

<210> 30
<211> 432
<212> PRT
<213> Phaffia rhodozyma ATCC 96594

<400> 30

Lys Glu Glu Ile Leu Val Ser Ala Pro Gly Lys Val Ile Leu Phe Gly
1 5 10 15

Glu His Ala Val Gly His Gly Val Thr Gly Ile Ala Ala Ser Val Asp
20 25 30

Leu Arg Cys Tyr Ala Leu Leu Ser Pro Thr Ala Thr Thr Thr Thr Ser
35 40 45

Ser Ser Leu Ser Ser Thr Asn Ile Thr Ile Ser Leu Thr Asp Leu Asn
50 55 60

Phe Thr Gln Ser Trp Pro Val Asp Ser Leu Pro Trp Ser Leu Ala Pro
65 70 75 80

Asp Trp Thr Glu Ala Ser Ile Pro Glu Ser Leu Cys Pro Thr Leu Leu
85 90 95

Ala Glu Ile Glu Arg Ile Ala Gly Gln Gly Gly Asn Gly Gly Glu Arg
100 105 110

```
Glu Lys Val Ala Thr Met Ala Phe Leu Tyr Leu Leu Val Leu Leu Ser
        115                 120                 125

Lys Gly Lys Pro Ser Glu Pro Phe Glu Leu Thr Ala Arg Ser Ala Leu
    130                 135                 140

Pro Met Gly Ala Gly Leu Gly Ser Ser Ala Ala Leu Ser Thr Ser Leu
145                 150                 155                 160

Ala Leu Val Phe Leu Leu His Phe Ser His Leu Ser Pro Thr Thr Thr
                165                 170                 175

Gly Arg Glu Ser Thr Ile Pro Thr Ala Asp Thr Glu Val Ile Asp Lys
            180                 185                 190

Trp Ala Phe Leu Ala Glu Lys Val Ile His Gly Asn Pro Ser Gly Ile
        195                 200                 205

Asp Asn Ala Val Ser Thr Arg Gly Gly Ala Val Ala Phe Lys Arg Lys
    210                 215                 220

Ile Glu Gly Lys Gln Glu Gly Gly Met Glu Ala Ile Lys Ser Phe Thr
225                 230                 235                 240

Ser Ile Arg Phe Leu Ile Thr Asp Ser Arg Ile Gly Arg Asp Thr Arg
                245                 250                 255

Ser Leu Val Ala Gly Val Asn Ala Arg Leu Ile Gln Glu Pro Glu Val
            260                 265                 270

Ile Val Pro Leu Leu Glu Ala Ile Gln Gln Ile Ala Asp Glu Ala Ile
        275                 280                 285

Arg Cys Leu Lys Asp Ser Glu Met Glu Arg Ala Val Met Ile Asp Arg
    290                 295                 300

Leu Gln Asn Leu Val Ser Glu Asn His Ala His Leu Ala Ala Leu Gly
305                 310                 315                 320

Val Ser His Pro Ser Leu Glu Glu Ile Ile Arg Ile Gly Ala Asp Lys
                325                 330                 335

Pro Phe Glu Leu Arg Thr Lys Leu Thr Gly Ala Gly Gly Gly Gly Cys
            340                 345                 350

Ala Val Thr Leu Val Pro Asp Asp Phe Ser Thr Glu Thr Leu Gln Ala
        355                 360                 365
```

Leu Met Glu Thr Leu Val Gln Ser Ser Phe Ala Pro Tyr Ile Ala Arg
370 375 380

Val Gly Gly Ser Gly Val Gly Phe Leu Ser Ser Thr Lys Ala Asp Pro
385 390 395 400

Glu Asp Gly Glu Asn Arg Leu Lys Asp Gly Leu Val Gly Thr Glu Ile
405 410 415

Asp Glu Leu Asp Arg Trp Ala Leu Lys Thr Gly Arg Trp Ser Phe Ala
420 425 430

<210> 31
<211> 4135
<212> DNA
<213> Phaffia rhodozyma ATCC96594

<400> 31
actgactcgg ctaccggaaa atatcttttc aggacgcctt gatcgttttg gacaacacca 60

tgatgtcacc atatcttcag cggccgttgg agctaggagt agacattgta tacgactctg 120

gaacaaagta tttgagtgga caccacgatc tcatggctgg tgtgattact actcgtactg 180

aggagattgg gaaggttcgt gcttgcttgc tttgaatgtc gtgcctaaag ccattgccat 240

aagacagagt ctgatctatg tcgtttgcct acaacagaga atggcctggt tcccaaatgc 300

tatgggaaat gcattgtctc cgttcgactc gttccttctt ctccgaggac tcaaaacact 360

tcctctccga ctggacaagc agcaggcctc atctcacctg atcgcctcgt acttacacac 420

cctcggcttt cttgttcact accccggtct gccttctgac cctgggtacg aacttcataa 480

ctctcaggcg agtggtgcag gtgccgtcat gagctttgag accggagata tcgcgttgag 540

tgaggccatc gtgggcggaa cccgagtttg gggaatcagt gtcagtttcg gagccgtgaa 600

cagtttgatc agcatgcctt gtctaatgag gttagttctt atgccttctt ttcgcgcctt 660

ctaaaatttc tggctgacta attgggtcgg tctttccgtt cttgcatttc agtcacgcat 720

ctattcctgc tcaccttcga gccgagcgag gtctccccga acatctgatt cgactgtgtg 780

tcggtattga ggaccctcac gatttgcttg atgatttgga ggcctctctt gtgaacgctg 840

gcgcaatccg atcagtctct acctcagatt catcccgacc gctcactcct cctgcctctg 900

attctgcctc ggacattcac tccaactggg ccgtcgaccg agccagacag ttcgagcgtg 960

ttaggccttc taactcgaca gccggcgtcg aaggacagct tgccgaactc aatgtagacg 1020

atgcagccag acttgcgggc gatgagagcc aaaaagaaga aattcttgtc agtgcaccgg 1080

gaaaggtcat tctgttcggc gaacatgctg taggccatgg tgttgtgagt gagaaatgaa 1140

agctttatgc tctcattgca tcttaacttt tcctcgcctt ttttgttctc ttcatcccgt 1200

cttgattgta gggatgcccc cctttgcccc tttccccttc ttgcatctgt ctatatttcc 1260

ttatacattt cgctcttaag agcgtctagt tgtaccttat aacaaccttt ggttttagca 1320

tcctttgatt attcatttct ctcatccttc ggtcagaggc tttcggccat ctttacgtct 1380

gattagattg taatagcaag aactatcttg ctaagccttt tctcttcctc ttcctcctat 1440

ataaatcgaa ttcactttcg gacatgttta ttttggggaa atcatcaagg ggtggggggc 1500

caatcccgac actaattttc tgctcacgtc aaaactcagc gttcagaatc agtcactgac 1560

cctgatacgt gtctctatgt gtgtgggtgt acgtgcgaat tgtgactcga cgttctacgc 1620

ttaaaaacag accgggatcg ctgcttccgt tgatcttcga tgctacgctc ttctctcacc 1680

cactgctacg acaacaacat catcgtcgtt atcgtctaca aacattacca tctccctaac 1740

ggacctgaac tttacgcagt cttggcctgt tgattctctt ccttggtcac ttgcgcctga 1800

ctggactgag gcgtctattc cagaatctct ctgcccgaca ttgctcgccg aaatcgaaag 1860

gatcgctggt caaggtggaa acggaggaga aagggagaag gtggcaacca tggcattctt 1920

gtatttgttg gtgctattga gcaaagggaa gccaaggtag gtttttttctg tctcttcttt 1980

ttgcctataa agactcttaa ctgacggaga aagtgttggg tttcttcctt cgggggttca 2040

atcaattaaa gtgagccgtt cgagttgacg gctcgatctg cgcttccgat gggagctggt 2100

ctgggttcat ccgccgctct atcgacctct cttgccctag tctttcttct ccacttttct 2160

cacctcagtc caacgacgac tggcagagaa tcaacaatcc cgacggccga cacagaagta 2220

attgacaaat gggcgttctt agctgaaaaa gtcatccatg gaaatccgag tgggattgat 2280

aacgcggtca gtacgagagg aggcgctgtt gctttcaaaa gaaagattga gggaaaacag 2340

gaaggtggaa tggaagcgat caagaggtac gcagacacgg tgcttcatat gccatactcc 2400

agtctgattg acccatgatg aacgtctttc tacatttcga atatagcttc acatccattc 2460

gattcctcat cacagattct cgtatcggaa gggatacaag atctctcgtt gcaggagtga 2520

atgctcgact gattcaggag ccagaggtga tcgtcccttt gttggaagcg attcagcaga 2580

ttgccgatga ggctattcga tgcttgaaag attcagagat ggaacgtgct gtcatgatcg 2640

atcgacttca agttagttct tgttcctttc aagactcttt gtgacattgt gtcttatcca 2700

tttcatcttc ttttttcttc cttcttctgc agaacttggt ctccgagaac cacgcacacc 2760

tagcagcact tggcgtgtcc cacccatccc tcgaagagat tatccggatc ggtgctgata 2820

agcctttcga gcttcgaaca aagttgacag gcgccggtgg aggtggttgc gctgtaaccc 2880

tggtgcccga tggtaaagtc tctccttttc tcttccgtcc aagcgacaca tctgaccgat 2940

gcgcatcctg tacttttggt caaccagact tctcgactga aacccttcaa gctcttatgg 3000

agacgctcgt tcaatcatcg ttcgcccctt atattgcccg agtgggtggt tcaggcgtcg 3060

gattcctttc atcaactaag gccgatccgg aagatgggga gaacagactt aaagatgggc 3120

tggtgggaac ggagattgat gagctagaca gatgggcttt gaaaacgggt cgttggtctt 3180

ttgcttgaac gaaagatagg aaacggtgat tagggtacag atcctttgct gtcatttta 3240

caaaacactt tcttatgtct tcatgactca acgtatgccc tcatctctat ccatagacag 3300

cacggtacct ctcaggtttc aatacgtaag cgttcatcga caaaacatgc ggcacacgaa 3360

aacgagtgga tataagggag aagagagata ttagagcgaa aaagagaaga gtgagagagg 3420

aaaaaaataa ccgagaacaa cttattccgg tttgttagaa tcgaagatcg agaaatatga 3480

agtacatagt ataaagtaaa gaagagaggt ttacctcaga ggtgtgtacg aaggtgagga 3540

caggtaagag gaataattga ctatcgaaaa aagagaactc aacagaagca ctgggataaa 3600

gcctagaatg taagtctcat cggtccgcga tgaaagagaa attgaaggaa gaaaaagccc 3660

ccagtaaaca atccaaccaa cctcttggac gattgcgaaa cacacacacg cacgcggaca 3720

tatttcgtac acaaggacgg gacattcttt ttttatatcc gggtggggag agagagggtt 3780

atagaggatg aatagcaagg ttgatgtttt gtaaaaggtt gcagaaaaag gaaagtgaga 3840

gtaggaacat gcattaaaaa cctgcccaaa gcgatttata tcgttcttct gttttcactt 3900

ctttccgggc gctttcttag accgcggtgg tgaagggtta ctcctgccaa ctagaagaag 3960

caacatgagt caaggattag atcatcacgt gtctcatttg acgggttgaa agatatattt 4020

agatactaac tgcttcccac gccgactgaa aagatgaatt gaatcatgtc gagtggcaac 4080

gaacgaaaga acaaatagta agaatgaatt actagaaaag acagaatgac tagaa 4135

<210> 32
<211> 1164
<212> DNA
<213> Artificial sequence

<220>
<223> Paracoccus zeaxanthinifaciens I17T

<400> 32
atgagaggat cgcatcacca tcaccatcac tcgaccggca ggcctgaagc aggcgcccat 60

gccccgggca agctgaccct gtccggggaa cattccgtgc tctatggtgc gcccgcgctt 120

gccatggcca tcgcccgcta taccgaggtg tggttcacgc cgcttggcat tggcgagggg 180

atacgcacga cattcgccaa tctctcgggc ggggcgacct attcgctgaa gctgctgtcg 240

gggttcaagt cgcggctgga ccgccggttc gagcagttcc tgaacggcga cctaaaggtg 300

cacaaggtcc tgacccatcc cgacgatctg gcggtctatg cgctggcgtc gcttctgcac 360

gacaagccgc cggggaccgc cgcgatgccg ggcatcggcg cgatgcacca cctgccgcga 420

ccgggtgagc tgggcagccg gacggagctg cccatcggcg cgggcatggg gtcgtctgcg 480

gccatcgtcg cggccaccac ggtcctgttc gagacgctgc tggaccggcc caagacgccc 540

gaacagcgct tcgaccgcgt ccgcttctgc gagcggttga agcacggcaa ggccggtccc 600

atcgacgcgg ccagcgtcgt gcgcggcggg cttgtccgcg tgggcgggaa cgggccgggt 660

tcgatcagca gcttcgattt gcccgaggat cacgaccttg tcgcgggacg cggctggtac 720

tgggtactgc acgggcgccc cgtcagcggg accggcgaat gcgtcagcgc ggtcgcggcg 780

gcgcatggtc gcgatgcggc gctgtgggac gccttcgcag tctgcacccg cgcgttggag 840

gccgcgctgc tgtctggggg cagccccgac gccgccatca ccgagaacca gcgcctgctg 900

gaacgcatcg gcgtcgtgcc ggcagcgacg caggccctcg tggcccagat cgaggaggcg 960

ggtggcgcgg ccaagatctg cggcgcaggt tccgtgcggg gcgatcacgg cggggcggtc 1020

ctcgtgcgga ttgacgacgc gcaggcgatg gcttcggtca tggcgcgcca tcccgacctc 1080

gactgggcgc ccctgcgcat gtcgcgcacg ggggcggcac ccggccccgc gccgcgtgcg 1140

caaccgctgc cggggcaggg ctga 1164

<210> 33
<211> 1164
<212> DNA
<213> Artificial sequence

<220>
<223> Paracoccus zeaxanthinifaciens I17T, G47D, K93E, P132S

<400> 33
atgagaggat cgcatcacca tcaccatcac tcgaccggca ggcctgaagc aggcgcccat 60

gccccgggca agctgaccct gtccggggaa cattccgtgc tctatggtgc gcccgcgctt 120

gccatggcca tcgcccgcta taccgaggtg tggttcacgc cgcttgacat tggcgagggg 180

atacgcacga cattcgccaa tctctcgggc ggggcgacct attcgctgaa gctgctgtcg 240

gggttcaagt cgcggctgga ccgccggttc gagcagttcc tgaacggcga cctaaaggtg 300

cacgaggtcc tgacccatcc cgacgatctg gcggtctatg cgctggcgtc gcttctgcac 360

gacaagccgc cggggaccgc cgcgatgccg ggcatcggcg cgatgcacca cctgccgcga 420

tccggtgagc tgggcagccg gacggagctg cccatcggcg cgggcatggg gtcgtctgcg 480

gccatcgtcg cggccaccac ggtcctgttc gagacgctgc tggaccggcc caagacgccc 540

gaacagcgct tcgaccgcgt ccgcttctgc gagcggttga agcacggcaa ggccggtccc 600

atcgacgcgg ccagcgtcgt gcgcggcggg cttgtccgcg tgggcgggaa cgggccgggt 660

tcgatcagca gcttcgattt gcccgaggat cacgaccttg tcgcgggacg cggctggtac 720

tgggtactgc acgggcgccc cgtcagcggg accggcgaat gcgtcagcgc ggtcgcggcg 780

gcgcatggtc gcgatgcggc gctgtgggac gccttcgcag tctgcacccg cgcgttggag 840

gccgcgctgc tgtctggggg cagccccgac gccgccatca ccgagaacca gcgcctgctg 900

gaacgcatcg gcgtcgtgcc ggcagcgacg caggccctcg tggcccagat cgaggaggcg 960

ggtggcgcgg ccaagatctg cggcgcaggt tccgtgcggg gcgatcacgg cggggcggtc 1020

ctcgtgcgga ttgacgacgc gcaggcgatg gcttcggtca tggcgcgcca tcccgacctc 1080

gactgggcgc ccctgcgcat gtcgcgcacg ggggcggcac ccggccccgc gccgcgtgcg 1140

caaccgctgc cggggcaggg ctga 1164